Received: 3 October 2019 | Revised: 23 September 2020 | Accepted: 12 October 2020

DOI: 10.1002/hec.4184

RESEARCH ARTICLE

Health Economics WILEY

Does commuting mode choice impact health?

Nikita Jacob[1] | Luke Munford[2] | Nigel Rice[3] | Jennifer Roberts[4]

[1]Centre for Health Economics, University of York, York, UK

[2]School of Health Sciences, University of Manchester, Manchester, UK

[3]Centre for Health Economics & Department of Economics and Related Studies, University of York, York, UK

[4]Department of Economics, University of Sheffield, Sheffield, UK

Correspondence
Nigel Rice, Centre for Health Economics & Department of Economics and Related Studies, University of York, York YO10 5DD, UK.
Email: nigel.rice@york.ac.uk

Funding information
Economic and Social Research Council, Grant/Award Number: ES/P007546/1; Medical Research Council Skills Development Fellowship, Grant/Award Number: MR/N015126/1

Abstract

Governments around the world are encouraging people to switch away from sedentary modes of travel towards more active modes, including walking and cycling. The aim of these schemes is to improve population health and to reduce emissions. There is considerable evidence on the latter, but relatively little on the former. This paper investigates the impact of mode choice on physical and mental health. Using data from the UK Household Longitudinal Study, we exploit changes in mode of commute to identify health outcome responses. Individuals who change modes are matched with those whose mode remains constant. Overall we find that mode switches affect both physical and mental health. When switching from car to active travel we see an increase in physical health for women and in mental health for both genders. In contrast, both men and women who switch from active travel to car are shown to experience a significant reduction in their physical health and health satisfaction, and a decline in their mental health when they change from active to public transport.

KEYWORDS
commuting mode, health, panel data econometrics

JEL CLASSIFICATION
C1, I1

1 | INTRODUCTION

Governments around the world are encouraging people to switch away from cars and toward more active modes of travel, including walking and cycling. For example, in the United Kingdom in 2017, the government announced a £1.2 bn scheme to encourage walking and cycling.[1] The aims of such schemes are twofold: (1) to improve population health by encouraging physical activity and (2) to reduce emissions and pollution levels. There is considerable evidence on the effectiveness of the latter aim (Rabl & deNazelle, 2012). There is also evidence on the benefits to health of more active modes of travel, and it is strongly recommended by the UK National Institute for Health and Care Excellence (NICE, 2012) as a feasible way of incorporating greater levels of physical activity into daily life. However, the majority of existing evidence has relied on cross-sectional data (Flint & Cummins, 2016; Flint, Cummins, & Sacker, 2014) which present challenges for causal interpretation. Although these cross-sectional associations are important, without further information they often suffer from endogeneity bias, due to the joint choices individuals may make over commuting mode and health.[2]

In order to estimate meaningful effects of the impact of mode of transport, we need to address issues of unobserved preferences and changes in mode choices occurring due to health related reasons. We tackle this by providing evidence of the effects of changes in commuting mode on health for adults in employment in the United Kingdom. Commuting is the most frequent reason for travel for working age individuals; the average UK commuter spends nearly an hour a day traveling and this is increasing over time (Department for Transport, 2017).

Taking advantage of the UK Household Longitudinal Study (UKHLS) which has a large sample size, and rich information on health and labor market experiences, we analyze the effect of changes to commuting mode on physical and mental health. A key feature of the data is that there are a sufficient number of individuals who are observed to change their mode of commute (the "treated" individuals) and we have an extensive pool of potential controls for whom commuting mode remains constant. We derive estimates of average treatment effects on the treated (ATTs) by exploiting matching methods (via entropy balancing [EB]; Hainmueller, 2012; Hainmueller & Xu, 2013). We are able to obtain a close balance on confounding covariates, that in part determine both health outcomes and commuting mode choice, across treated and control individuals. Following Ho, Imai, King, and Stuart (2007), we do this by preprocessing the data via matching prior to undertaking parametric modeling. This "doubly robust" approach has the advantage of being robust to either misspecification in the parametric model but complete covariate balance via matching, or incomplete balance through matching but correct specification of the regression model. This can be viewed as a way to achieve balance in covariates with the objective of reducing model dependence in the subsequent regressions to extract the ATTs (Abadie & Imbens, 2011).

We follow individuals over time until they change their mode of commute, and compare their health responses to those of a matched control group. We match on socio-demographic characteristics observed pre-treatment, including initial mode, duration of commute and health status. Using regression methods we then compare health outcomes between treated individuals and their matched controls. Conditional on the validity of selection on observables[3], the approach identifies a causal effect of a change in commuting mode on health. Our main outcomes of interest are summary measures of mental and physical health derived from the Short Form 12 (SF12), and self-reported satisfaction with health. Our findings show that adopting active means of travel improves health, for both men and women. Changing from an active mode to either public transport or car travel has a negative impact on health. Further analyses, comparing outcomes in the short and intermediate term, confirm our main results.

2 | RELATED LITERATURE

A number of studies have looked at the relationship between mode of commuting and health/well-being. The general consensus is that active commuting has positive effects on physical, mental and overall general health. Evidence from the United Kingdom has consistently suggested that levels of physical activity involved in active modes of commuting, such as walking or cycling, translate into health benefits for individuals; including lower BMI and body fat, enhanced mood and increases in mental and physical health.[4] For example, Laverty, Mindell, Webb, and Millett (2013), using data from the first wave of the UKHLS, show that, in comparison to the use of private means of transport, the use of public transport, as well as walking or cycling to work, was associated with a lower likelihood of being overweight. Individuals who walked or cycled to work had a lower likelihood of having diabetes, and individuals who walked had a lower likelihood of having hypertension.

The mental health benefits of active travel arise from the fact that it is perceived to be both more relaxing and exciting than other modes of transport (Scheepers et al., 2014). It also promotes higher life satisfaction (Morris, 2015) and is associated with a lower rate of mental distress. MacDonald, Stokes, Cohen, Kofner, and Ridgeway (2010) and Frank, Andresen, and Schmid (2004) suggest that spending more time in cars is associated with increases in obesity and blood pressure, perhaps due to the frustrations of traffic congestion (Stokols, Novaco, Stokols, & Campbell, 1978). Other studies have also concluded that car commuting is stressful and leads to negative mood among drivers[5] Contrasting evidence, however, by Anable and Gatersleben (2005) and Eriksson, Friman, and Gärling (2013) has shown that driving to work provides individuals a positive feeling through greater control and flexibility over their commute. Active travel also has positive effects on the environment since it reduces air pollution (Rabl & deNazelle, 2012), which in turn reduces the risk of cancer (Litman, 2010) and cardiovascular diseases (Genter, Donovan, Petrenas, & Badland, 2008; Hamer & Chida, 2008; Litman, 2010; Scheepers et al., 2014).

Studies outside the UK report similar evidence. Turcotte (2005, 2011), and Páez and Whalen (2010) using Canadian data and Friman, Fujii, Ettema, Grling, and Olsson (2013) using Swedish data, find that active travel commuters tend to report higher satisfaction than users of other modes; public transport users were least satisfied.[6] However, in terms of the effects on health, several studies have concluded that public transport users tend to be physically healthier than car commuters since they meet the recommended level of physical activity more often, as they tend to walk to reach bus or train terminals (MacDonald et al., 2010; Wener & Evans, 2007). Other studies suggest that using public transport causes travelers to experience lower levels of stress (Wener & Evans, 2011).

Little research has explored the effects of *changes* in travel mode on health. Martin, Goryakin, and Suhrcke (2014) explore the relationship between active travel and psychological well-being using British Household Panel Survey data from 1991–2009. The study relies on fixed effects models to investigate how choice of travel mode, commuting time and switching to active travel impact psychological well-being. They found evidence to suggest that switching to active travel was associated with an improvement in well-being compared to individuals who always commuted by car or public transport. Extending their study using the same dataset, Martin, Panter, Suhrcke, and Ogilvie (2015) examined the effect of switching from private motor transport to active travel or public transport (in the next period) on changes in BMI. They found that those who switched saw a reduction in BMI within a 2 year period.

We advance this literature by taking into account the potential for selection bias and exploiting methods of matching together with parametric regression, to improve identification of the health impacts of commuting mode choice. We only consider individuals for whom household location is fixed but allow job locations to vary; which may be employer or employee induced.[7] A change in job location may lead to a change in commuting mode through either a change in commuting route and/or distance, or a change in job remuneration allowing, via an income effect, greater choice of travel mode.

3 | CONCEPTUAL FRAMEWORK

We assume that individuals derive utility (or disutility) from commuting, such that $U = U(m, h(m, t), z)$, where m represents mode choice and z represents other consumption from which individuals derive utility. Individuals are also assumed to value any health impacts of their commuting mode choice, which will also be a function of time spent commuting represented by $h(m, t)$. Hence individuals derive utility, both directly and indirectly, through their choice of commuting mode. Direct utility may be positive, for example, the enjoyment of driving, the ability to relax or work on public transport, the enjoyment of exercise from walking or cycling to work, or negative, for example, frustration of sitting in heavy traffic, crowded public transport, inclement weather during active commuting. Indirect utility is derived from mode choice through the impact this has on health (Frank et al., 2004; Lancee, Veenhoven, & Burger, 2017; MacDonald et al., 2010; Wener & Evans, 2007). For example, exposure to exhaust fumes or being seated for long periods of time might impact physical health; the uncertainty of disruption during car travel may affect mental wellbeing. Accordingly, commuting mode can be seen as being valued for both a consumption property—the direct impact on utility, and an investment property—the indirect health effects (Grossman, 1972). In making choices over mode, individuals are assumed to maximize utility subject to constraints over income and time. Different forms of travel attract different prices and hence cost to the commuter and are therefore influenced by an individual's income constraint. Individuals also face a time constraint, which, during the working day, consists of choices over time spent on leisure (t_l), work hours (t_w), and commuting (t_c), such that $(t_l + t_w + t_c = 24 \text{ hrs})$. The greater time spent commuting, the less time available for other pursuits, assumed mainly to be leisure for individuals with fixed hours of work. In this way, commuting entails an opportunity time cost to the individual and choices over mode will be influenced by this constraint. Individuals are assumed to choose the commuting mode that maximizes their utility subject to the constraints they face at a particular point in time. Should the value individuals place on the investment and/or consumption properties of mode choice change, or should individuals face changes to their constraints (e.g., through a change in job location or road infrastructure), this may lead to a change in commuting mode.

We are interested in identifying the health effects of commuting mode choice. Our approach considers those individuals who change mode at time t as treated and those who do not change mode as potential controls. By matching controls to treated individuals at time $t-1$ we assume that the average utility of the two groups, prior to treatment, is equivalent. Matching is undertaken on a set of potential confounding characteristics thought, a priori, to influence both mode choice and health; this includes initial mode and commuting time, health status, and household income among other factors.[8]

Adopting a potential outcomes framework, the above procedure assumes that conditional on the set of confounding covariates, x, selection into treatment, d, is independent of potential outcomes, such that $(h^0, h^1) \perp d|x$. where h^0 and h^1 are potential health outcomes for treated individuals without treatment, h^0, and with treatment, h^1, respectively. This is often termed the conditional independence assumption (Heckman and Robb (1985)). Where this holds, the $ATT = E(h^1 - h^0 \mid x, d=1) = E(h^1 \mid x, d=1) - E(h^0 \mid x, d=1)$ can be estimated by replacing the unobserved component $E(h^0 \mid x, d=1)$ with its observed counterfactual $E(h^0 \mid x, d=0)$. The conditional independence assumption is required to hold for us to identify the causal effect of change in mode on health outcomes. This requirement is challenging in any empirical application that relies on techniques based on selection on observables (this is also true of methods that are based on selection on unobservables). However, we emphasize that we match on a wealth of pre-treatment characteristics, and importantly include pre-treatment mode choice and health. These characteristics are also used in the regression analysis which follows EB. By conditioning on prior health we are effectively considering changes to health due to a change in mode, mimicking a fixed effects approach. We further balance on variables that represent the sequence of observations observed for individuals in the panel dataset. This is intended to balance for potential attrition bias, which, in part, is likely to be driven by unobservables. Following matching, we estimate the treatment effect using a regression framework. The latter helps to mitigate bias resulting from less than perfect matching.

4 | EMPIRICAL APPROACH

Our empirical strategy exploits *changes* to mode of commute observed in the data at time t, but occurring somewhere between $t-1$ and t, to identify the responses on health outcomes at time $t+1$. We compare outcomes for 'treated' individuals who experience a change to their mode of commute with outcomes for observationally identical (as of $t-1$) controls, who do not experience a change to their commuting mode.[9] Prior to the occurrence of the change, observational equivalence is defined by a wide set of potential confounding variables, including demographic and individual factors such as age and sex, baseline health and educational achievement; household characteristics such as cohabiting status, number of kids, and household income; and labor market characteristics such as job hours and baseline mode of commuting. These variables are expected to determine, in part, both household health and choices over mode of transport. Conditioning on baseline health and commuting mode is important in defining suitable controls for individuals who are observed to change commuting mode. Including baseline health in both the matching and subsequent regression has the further advantage of identifying health effects from a change to commuting conditional on individual-specific underlying level of health. This helps to remove unobserved effects specific to individuals and their health.

Our approach follows the principles set out in Ho et al. (2007) to use matching methods to preprocess the data prior to parametric modeling of outcomes.[10] The aim is to reduce model dependency by using matching to create balance in covariate distributions across treated and control groups. Successful (perfect) matching renders treatment independent of control variables. Subsequent parametric regression modeling of the preprocessed data is therefore less dependent on specification assumptions and hence more likely to identify consistent causal effects. Where matching proves to be less than perfect, the application of regression techniques conditional on the same set of confounding variables controls for the lack of perfect balance. The approach can be viewed as an extension of the usual matching techniques, which rely on comparisons of means of the matched data.

The matching method we propose, together with subsequent regression modeling assumes that selection into a change of commuting mode can be captured by the set of conditioning variables used. If this assumption does not hold and selection is also a function of unobservable characteristics, then techniques such as instrumental variables would be required.[11] As with many empirical applications, it is profoundly difficult to find appropriate instruments to identify causal effects; and even where these can be found identification often leads to very localized treatment effects which suffer from a lack of generalizability. An important consideration when selecting variables to match controls to treated individuals is that these are based on observed characteristics measured prior to treatment. By doing so, we eliminate the possibility of treatment influencing the set of conditioning variables. As well as the set of characteristics used in previous studies, we further condition on initial commuting mode and health. In this way, the model mimics a fixed effects approach to dealing with individual-specific unobservable characteristics thought to influence outcomes and mode choice.

We match individuals based on their characteristics measured at $t-1$, which can lie anywhere between the first observed wave and the antepenultimate wave (for our main results outcomes are measured at time $t+1$). Sample attrition bias in panel data might arise due to healthier individuals remaining in the panel longer than less healthy counterparts. Health, at least in part, is assumed to determine and be determined by transport mode choice. To mitigate concerns that attrition bias may arise differentially across treated and control individuals we further match on the wave of mode change together with two variables constructed to better approximate the panel profile of treated and control individuals across waves of the data. These are informed by the literature on testing for attrition bias (Jones, Koolman, & Rice, 2006; Verbeek & Nijman, 1992). For each individual we construct variables to represent the total number of waves and the number of consecutive waves they are observed in the dataset.[12] These are included in the matching step.

Matching is undertaken for each of the observed treatments defined by changes in commuting mode: car-public, car-active, public-active and their converse. We then regress outcomes on the set of controls and a treatment effect separately for each of the six matched samples as follows:

$$h_{i,t+1} = \alpha + \beta_d d_{i,t} + X'_{i,t-1}\beta_x + \gamma\lambda_{i,t-1} + \epsilon_{i,t+1} \tag{1}$$

where β_d identifies the treatment effect of interest; the change in mode at time t on health outcomes, h_i at time $t+1$. The set of variables used to match controls to treated individuals prior to treatment are represented by $X_{i,t-1}$ (see Table 5 for the variables) and their corresponding relationship with outcomes, β_x.[13] $\lambda_{i,t-1}$ are wave indicators to recognize that mode changes may occur in different calendar years; $\epsilon_{i,t+1}$ is the usual idiosyncratic error term. Regression weights derived from EB are applied to Model (1). Models for cardinal outcomes are estimated using ordinary least squares; ordered categorical outcomes are estimated with ordered probits. All regressions contain robust standard errors.

We use matching techniques to adjust the covariate distribution of the control group data by reweighting and/or discarding units such that it becomes more similar to the covariate distribution in the treatment group. We apply EB (Hainmueller, 2012), which involves a reweighting scheme that directly incorporates covariate balance into the weight function that is applied to the sample units. This is done by selecting a set of weights for each observation in the control group that minimize an entropy distance metric subject to balance and normalizing constraints. This ensures that the weights are nonnegative and sum to unity. These weights satisfy a set of balancing constraints that involve specifying exact balance on moments of the covariate distributions (in our case the mean and variance) in the treatment and the reweighted control group.

All individuals are considered untreated in the first wave. An individual is assigned only once to the treatment group, when they first change their mode of commute, any subsequent changes in commuting mode are excluded from analysis.[14] Treated individuals never act as potential controls at any other point in time. Potential control individuals are those who never change their mode of commute while they are observed.

We are concerned with three different commuting modes; car, public transport and active travel; and consider the following changes: car to active travel, public transport to active travel, active travel to car, and active travel to public transport. We have additionally considered switches between car and public transport, but as these do not involve a switch into or out of more active modes, which are often the policy goal, these are not the main focus of our analysis. For each change in mode we match control individuals to treated individuals and then perform regression analysis on the balanced data. Matching is undertaken at $t-1$, mode change is observed at time t and outcomes at $t+1$. We further repeat the analyses (including matching and regression on outcomes) to compare short-run outcomes at time t, and longer term outcomes at $t+2$.[15]

An important feature of the literature on commuting is the difference in travel behavior between men and women, with men, on average, undertaking longer commutes. Further, Roberts, Hodgson, and Dolan (2011), find that the wellbeing of women, but not men, is adversely affected by increased commuting times, while Jacob, Munford, Rice, and Roberts (2019) provide evidence that this is due to the different labor markets in which women and men operate. Accordingly, we undertake heterogeneity analysis by gender and apply EB and regression analysis within gender for each of the mode changes.

5 | DATA

5.1 | UK Household Longitudinal Study

The UKHLS is a nationally representative sample of UK households, containing panel information on around 100,000 individuals in 40,000 households. We use seven waves of data from 2009 to 2016, containing rich information on

TABLE 1 Information on inclusion criteria and sample size

Criteria	Number		Percent	
	Observations	Individuals	Observations	Individuals
	NT	*N*	*NT*	*N*
Full UK Household Longitudinal Study sample	333,773	83,287	100%	100%
In at least two waves	315,330	64,844	94%	78%
Employed in all waves	148,218	38,365	44%	46%
No change of house	127,030	35,908	38%	43%
Non-missing work travel information	119,243	33,620	36%	40%
Non-missing health indicators	108,292	32,247	32%	39%
Age ≥ 16 and ≤ 65	106,464	31,787	32%	38%
Non-missing education, job hours, other health information	106,195	31,736	32%	38%
Surveyed for $\geq$ 3 waves	86,519	18,156	26%	22%
Surveyed for $\geq$ 4 waves	73,715	13,888	22%	17%

socio-economic, health, and labor market characteristics. Health is measured using component scores derived from the SF12 questionnaire. The SF12 uses twelve questions to measure functional health and wellbeing; the responses are aggegated to form the Physical (SF12-PCS) and Mental (SF12-MCS) Component Scores. These are cardinal representations of underlying health status, designed to lie between 0 (lowest level of health) and 100 (highest), and have a mean of 50 and a standard deviation of 10 for the general population (Ware, Keller, & Kosinski, 2002). As an additional outcome we also use responses to questions on satisfaction with health, recorded on a five point ordered categorical scale, where 1 is least satisfied and 5 is most satisfied.[16]

Our measure of commuting mode is taken from the question "*How do you usually get to your place of work?*" which is asked only to people who state they are in employment. Responses are categorized as Car (drivers and passengers), Public transport (bus, train, undergound) and Active travel (cycle, walk) with Other (taxi, moped, other mode) as an alternative group that we do not consider due to small sample sizes. To control for individual preferences we condition on characteristics typically used in the literature, including age, educational attainment, the number of children in a household, a married/cohabiting identifier, and log equivalised monthly household income (deflated to 2005 prices, and equivalised using the OECD modified scale, detailed in Foster, 2009).

Table 1 presents information on the basic inclusion criteria for the estimation sample. The seven waves of the UKHLS contain information on $N = 83,287$ individuals who are observed across waves to provide $NT = 333,773$ observations. We remove individuals who are observed in only a single wave; individuals not employed and individuals who change place of residence. The criterion of being observed in at least two consecutive waves allows us to consider short-run outcomes at time t following balancing on covariates at time $t-1$. Our working age (16–65 years) sample consists of 31,736 individuals for whom there are a total of 106,195 observations.[17] Descriptive statistics for this sample are provided in Table 2. The mean scores on SF12 PCS (physical health) and SF12 MCS (mental health) are 52.9 and 49.9, respectively, while the mean for health satisfaction 3.5. There are slightly more observations on females than males; mean age is 42 years; 45% have a university level qualification, average usual hours of work is 33; and average log equivalized monthly household income is £7.55 (equivalent to £1900/month.) These figures are in line with average values of the UK workforce obtained from the 2011 Census and estimates from the UK Labour Force Survey.

First, the data are stratified into treated and control groups, where the treated are observed to change mode, for example, from car to active travel and the control group never change. Secondly, for this sub-sample, matching controls to treated individuals through EB is undertaken followed by weighted regression of outcomes (here at time t). Exact sample sizes will vary across the four possible mode changes observed. Our main outcome of interest is observed at time $t + 1$. Similarly, when considering long-run effects ($t + 2$), the initial basic sample is further refined to exclude individuals with less than four waves of data before matching and regression analysis.

TABLE 2 Summary statistics for estimation sample

	Overall					Women			Men		
	Mean	Std. Dev.	Min	Max	*N*	Mean	Std. Dev.	N	Mean	Std. Dev.	*N*
SF12PCS	52.901	8.035	4.64	74.710	106,195	52.773	8.439	58,927	53.059	7.497	47,268
SF12MCS	49.943	8.918	0	77.09	106,195	49.123	9.257	58,927	50.965	8.363	47,268
Satisfaction with own health	3.495	1.047	1	5	105,952	3.478	1.065	58,783	3.683	0.939	47,268
Male	0.445	0.497	0	1	106,195	-	-	-	-	-	-
Age	42.072	11.734	16	65	106,195	42.196	11.609	58,927	41.918	11.886	47,268
University level qualification	0.449	0.497	0	1	92,592	0.469	0.499	51,365	0.423	0.494	41,227
College level qualification	0.212	0.409	0	1	92,592	0.197	0.398	51,365	0.231	0.421	41,227
School level qualification	0.201	0.401	0	1	92,592	0.208	0.406	51,365	0.194	0.395	41,227
Household size	3.092	1.345	1	16	106,195	3.049	1.294	58,927	3.146	1.404	47,268
Number of children	0.707	0.979	0	8	106,195	0.681	0.939	58,927	0.739	1.025	47,268
Married/Cohabiting	0.712	0.453	0	1	106,055	0.682	0.466	58,837	0.749	0.434	47,218
Usual hours worked	33.186	10.334	0.1	97.7	106,195	29.614	10.247	58,927	37.638	8.561	47,268
Log household income	7.55	0.537	1.901	9.903	105,986	7.522	0.546	58,777	7.584	0.525	47,209

Note: Our working sample is $NT = 106{,}195$, based on an unbalanced sample of $N = 31{,}736$ individuals.

TABLE 3 Sample commuting times by gender and mode

	NT	Mean	Std. Dev.	Median
All modes				
Commuting time[a]—full sample	106,195	25.50	20.48	20
Male	47,268	27.83	22.11	20
Female	58,927	23.62	18.86	20
By mode[b]				
Car—all	74,181	23.19	17.81	20
Male	33,120	25.36	19.70	20
Female	41,061	21.43	15.92	20
Public transport—all	14,576	47.88	24.21	45
Male	6579	50.79	24.89	45
Female	7997	45.49	23.37	45
Walk or cycle—all	15,643	15.94	12.60	15
Male	6402	17.86	14.07	15
Female	9241	14.61	11.28	10

[a]We winsorize the commuting data to omit unrealistic extreme values. Observations above the 99th centile are recoded to be equal to the value at the 99th centile. This does not affect our conclusions.

[b]Car is defined as any commuter who uses either a car or van (driver or a passenger). Public transport is defined as those who use either a bus, train, or underground/tram. Note that the sum of Car + Public Transport + Walk or Cycle is not equal to the overall sample size as we do not include people who use a motorcycle, moped or taxi.

Table 3 breaks down the descriptive statistics of commuting time by gender and mode of transport. Males, in general, experience longer commutes (27.8 min one-way compared to 23.6 for women), with the differential between men and women remaining irrespective of the mode of transport. Public transport is associated with the longest commuting times (an average of 48 min) and cycling the shortest (16 min). The distribution of commuting times for

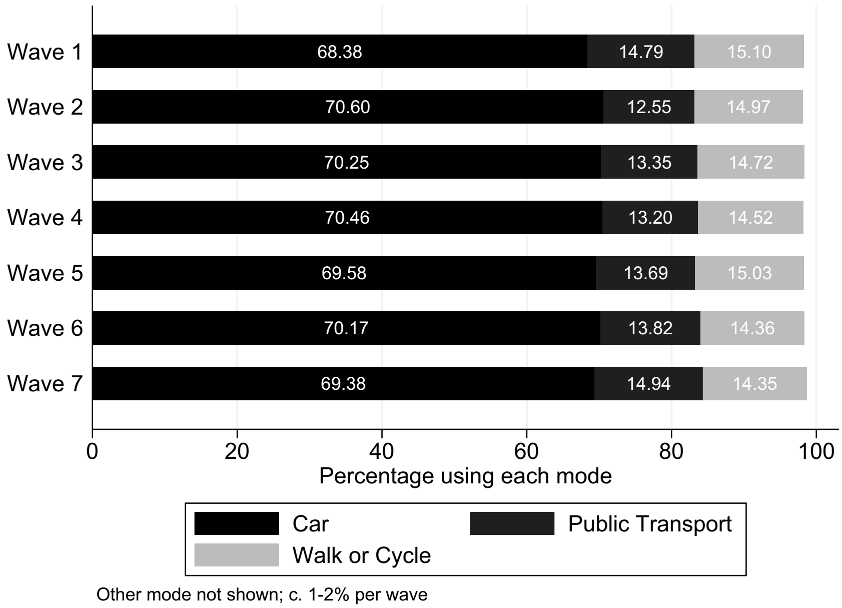

FIGURE 1 Percentage of individuals using each mode across all seven waves

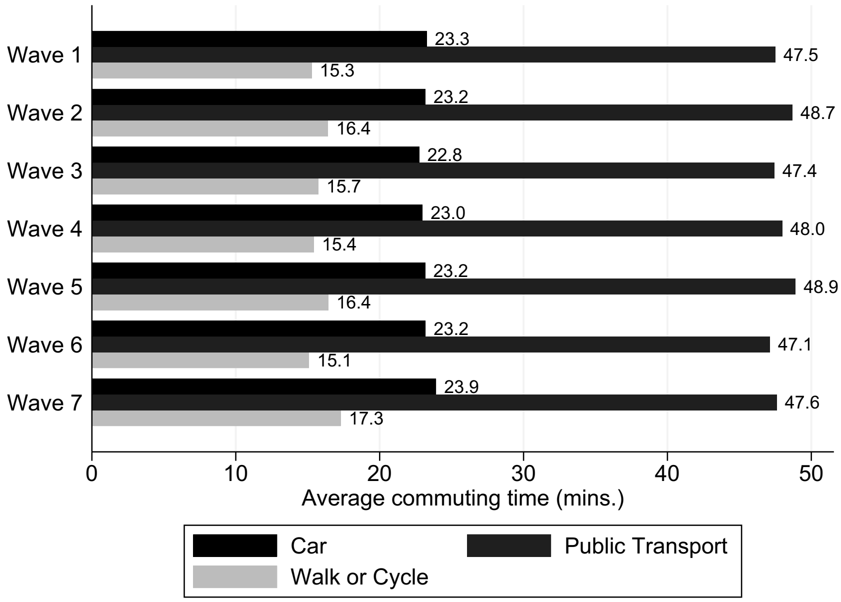

FIGURE 2 Average commuting time by mode of travel across all seven waves

active travel and non-active (users of public transport or car) is provided in the Appendix as Figure A1. As expected there is a greater concentration of short commute durations for active commuters compared to non-active.

Figure 1 shows the percentage of individuals who use each of the three modes over time. The percentage of people using a car is relatively stable at around 70% in each wave. The percentage using public transport drops between waves 1 and 2, but then steadily increases. There has been a slight decline in the number of people walking or cycling. Figure 2 shows the associated commuting times. All three modes have experienced a gradual increase in commuting time, but this is largest for walking and cycling.

Table 4 reports the transition probabilities between waves t and $t+1$. Among car users at time t, 95% will remain so in the following wave, with 2% switching to public transport and 3% to walking or cycling. Amongst initial public transport users, 81% remain whereas 13% switch to car and 6% switch to active modes. Finally, among initial active commuters, 78% remain so, whereas 16% and 5% switch to car and public transport, respectively. So in summary, there is much more resilience to switching away from car than the other two modes.

TABLE 4 Transition probabilities between waves

	$t+1$		
Mode	Car	Public transport	Walk or cycle
Car	0.95	0.02	0.03
Public transport	0.13	0.81	0.06
Walk or cycle	0.16	0.05	0.78

TABLE 5 Mode change: Car to public, entropy balancing match estimates

		Treated $N = 646$			Control $N = 39{,}636$		
		Mean	Variance	Skewness	Mean	Variance	Skewness
Before Balancing	Age	40.10	126.40	−0.14	43.44	109.40	−0.22
	Number of kids	0.75	1.02	1.19	0.75	0.97	1.15
	Job hours	33.61	94.24	−0.80	34.00	89.41	−0.48
	Married	0.67	0.22	−0.75	0.77	0.18	−1.29
	Household income (log)	7.55	0.34	0.10	7.57	0.25	−0.14
	SF12_PCS	53.49	64.79	−1.66	53.10	61.96	−1.65
	SF12_MCS	49.44	77.44	−1.06	50.19	74.61	−1.18
	CT_5 min (log)[c]	3.23	0.49	−0.19	2.93	0.51	−0.17
	Wave	2.92	2.38	0.39	3.41	2.57	0.08
	Treated wave	4.33	2.51	0.15	4.58	2.57	0.00
	Consecutive waves	3.26	2.77	0.18	3.81	2.50	−0.17
	Number of waves	4.88	1.76	0.07	5.28	1.65	−0.29
After balancing	Age	40.10	126.40	−0.14	40.44	125.90	−0.01
	Number of kids	0.75	1.02	1.19	0.75	1.01	1.24
	Job hours	33.61	94.24	−0.80	33.71	93.81	−0.52
	Married	0.67	0.22	−0.75	0.68	0.22	−0.79
	Household income (log)	7.55	0.34	0.10	7.54	0.34	−1.23
	SF12_PCS	53.49	64.79	−1.66	53.42	64.58	−1.81
	SF12_MCS	49.44	77.44	−1.06	49.51	77	−1.11
	CT_5 min (log)	3.23	0.49	−0.19	3.20	0.51	−0.35
	Wave	2.92	2.38	0.39	2.98	2.45	0.38
	Treated wave	4.33	2.51	0.15	4.36	2.53	0.15
	Consecutive waves	3.26	2.77	0.18	3.32	2.79	0.09
	Number of waves	4.88	1.76	0.07	4.93	1.77	0.01

Note: Matching using entropy balancing on first and second moments of covariate distribution. Except for treated wave, all are lagged (1) variables. Dependent variable measured at $t + 1$. Sample consists of individuals who are in the survey for at least three waves.

[c]Log of commuting time in 5 min bins. We account for attrition bias using the variables, *consecutive* waves and *number of waves*. Consecutive waves measures the number of times an individual is surveyed consecutively. Number of waves measures the total number of occurrences an individual makes in the sample.

6 | RESULTS

The success of any matching strategy is achieved through obtaining close covariate balance and common support between treated and controls. This relies on the availability of an adequate number of potential control individuals. From our sample, 82% (26,177) of individuals report no change in their commuting mode (the controls), while 12%

(3654) report having changed mode once across the sample period. The remaining observations are observed to change mode twice (5%) or more. A full breakdown is provided in Table A1.

Table 5 illustrates EB, on the first and second moments (mean and variance), for a mode change from car to public transport. Matching takes place on covariates measured at time $t-1$. Treated individuals undergo change in mode, controls remain as car users. EB equates the moments of the covariate distribution across treated and control groups. As can be seen, following EB the mean and variance of the set of covariates are very similar across treated and control individuals. This is reassuring as it provides support that the conditional independence assumption, $(h^0, h^1) \perp d|x$., set out in Section 3 holds. EB for other mode changes and for men and women separately (not reported here), follow a similar pattern.

The results in Table 6 exploit changes to commuting mode occurring between $t-1$ and t to identify health outcomes observed at $t + 1$. EB is used to preprocess the data using information on the set of controls prior to parametric modeling. The results suggest that mode changes from car to public transport and vice versa, do not impact health outcomes. Estimated effects are generally small and do not attain statistical significance. In contrast, when considering a mode change from car to active travel, we observe a large positive effect on mental health (SF12-MCS). The effect is observed, in similar magnitude, for both men and women. There is also an indication that physical health (SF12-PCS) improves for women, significant at the 10% level. Interestingly, individuals who switch mode from active to car report a significant decrease in physical health. Again these effects are observed overall and for men and women separately. We also observe a decrease in satisfaction with health for the overall sample (at 10% significance). It would appear, therefore, that the effect of a change from car to active travel is felt more strongly through improvements to mental health, whereas the effect of a change to car from active travel is felt through decreases to physical health. We speculate that the asymmetry might be due to the immediacy of feeling the effects. Improvements in mental health brought about through exercise are likely to be felt more quickly (e.g., through increased adrenaline and release of endorphins). One might expect this to be apparent when switching from car to regular active travel. A converse change in mode, however, might not produce such immediate effects. When we stop exercising while we might initially miss the high that this produced, over time we are more likely to pay greater attention to a feeling of lethargy and weight gain, and assign this to a decrease in physical, rather than mental, health. However, we do not observe the same effects when considering changes from public transport to active travel and vice-versa. Individuals who switch from public to active forms of travel report increased health satisfaction, predominantly men, but we do not observe significant effects for mental or physical health; however, this may be due to small sample sizes. The reverse mode change from active travel to public transport is associated with a reduction in mental health, particularly for men.[18] The contrast in results from a switch from active travel to car (decrease in physical health) compared to the switch from active travel to public transport (decrease in mental health) suggests that the experience of car and public transport confer different effects on the commuter. A user of public transport is passive, has no control over the journey and often is subjected to overcrowding. In contrast car users have some control over the journey and are actively engaged in the commuting process. The lack of control and passive nature of public transport may lead to a more noticeable effect on mental health. As a consequence, a switch from active travel to car is more likely to be experienced as a decrease in physical health. A graphical illustration of these results, for switches to or from active travel across both men and women is shown in Figure A2.

Overall, we do not observe effects on health from changes in mode between public transport and car use, or vice-versa, but do observe effects when moving between active and other forms of travel. However, effects appear generally small, typically less than a tenth of a standard deviation. In comparison to other studies that use the SF12 health measure, Ziebarth (2010) shows that the difference in means for the mental health score of the SF12 is 6.2 and physical health is 3.6 (when rescaled between 0 and 1), when comparing health for the lowest income percentile group to the highest percentile group. While the study does not explicitly consider changes in income and instead compares means across groups, the results do provide context to the size of effects found in this paper for observed changes in commuting mode. In general, our findings indicate that changing commuting mode has a notable impact on health. A change of mode from car to active travel for women has an approximate equivalent effect on physical health of one sixth of the effect of moving between the lowest and highest income percentile groups. The corresponding effect on mental health for both men and women is approximately equivalent to one eighth of the effect of changing income percentile groups.

6.1 | Immediate and longer run effects

Here we investigate the possible immediate effects (at time t), as well as longer-run effects (at time $t + 2$), of a change in commuting mode (occurring at time t). Full results are reported in Tables A2 and A3. Results are broadly

TABLE 6 Entropy Balancing by gender for outcomes at time $t + 1$

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health.Sat	Women (8) Health.Sat	Men (9) Health.Sat
Car->Public	0.037	0.047	0.142	0.000	−0.138	0.151	0.034	0.002	0.093
	(0.351)	(0.492)	(0.490)	(0.391)	(0.538)	(0.564)	(0.056)	(0.075)	(0.084)
N	29,714	16,954	12,760	29,714	16,954	12,760	29,617	16,895	12,722
Treated	646	369	277	646	369	277	644	367	277
Control	39,636	22,518	17,118	39,636	22,518	17,118	39,544	22,463	17,081
Car->Active	0.279	0.651*	−0.103	0.874***	0.821**	0.725*	0.050	0.060	0.038
	(0.250)	(0.344)	(0.363)	(0.290)	(0.413)	(0.396)	(0.042)	(0.056)	(0.062)
N	29,937	17,064	12,873	29,937	17,064	12,873	29,839	17,005	12,834
Treated	909	498	411	909	498	411	906	496	410
Control	39,636	22,518	17,118	39,636	22,518	17,118	39,544	22,463	17,081
Public->Car	−0.072	0.422	−0.691	0.347	0.593	0.108	0.024	0.044	0.035
	(0.358)	(0.513)	(0.487)	(0.436)	(0.623)	(0.610)	(0.059)	(0.077)	(0.091)
N	3909	2094	1815	3909	2094	1815	3889	2083	1806
Treated	707	412	295	707	412	295	706	411	295
Control	4639	2473	2166	4639	2473	2166	4619	2461	2158
Public->Active	−0.026	−0.805	0.947	0.306	0.769	−0.313	0.239***	0.146	0.358***
	(0.564)	(0.838)	(0.671)	(0.628)	(0.845)	(0.957)	(0.083)	(0.108)	(0.127)
N	3609	1912	1697	3609	1912	1697	3589	1901	1688
Treated	330	188	142	330	188	142	329	187	142
Control	4639	2473	2166	4639	2473	2166	4619	2461	2158
Active->Car	−0.824***	−0.815*	−0.752*	−0.165	−0.344	0.031	−0.077	−0.096	−0.011
	(0.302)	(0.428)	(0.407)	(0.369)	(0.486)	(0.594)	(0.048)	(0.065)	(0.075)
N	4098	2542	1556	4098	2542	1556	4083	2533	1550
Treated	861	487	374	861	487	374	856	483	373
Control	4688	2964	1724	4688	2964	1724	4678	2959	1719
Active->Public	−0.690	−0.917	−0.440	−1.010	0.137	−2.597***	−0.065	−0.033	−0.117
	(0.468)	(0.673)	(0.699)	(0.673)	(0.943)	(0.974)	(0.084)	(0.119)	(0.122)
N	3670	2305	1365	3670	2305	1365	3659	2299	1360
Treated	333	196	137	333	196	137	332	195	137
Control	4688	2964	1724	4688	2964	1724	4678	2959	1719

Note: Individuals present for at least 3 waves. Dependent variables measured at $t + 1$, and are increasing in good health. Controls matched to treated using entropy balancing at $t{-}1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t{-}1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

similar to those reported in Table 6. Mode changes from car to public transport or vice-versa, do not lead to changes in reported health or health satisfaction (an exception is that women report increased health satisfaction from a change from public to car at time $t + 1$). A shorter run effect of a change from car to active travel is observed for women's mental health and for health satisfaction overall. We also observe a decrease in reported physical health for men in the short-run when switching from active travel to car. These results echo those observed for the main results at time $t + 1$. We also observe shorter-run effects of a reduction in health satisfaction (at $p < 10\%$ level). Similarly, mode change from public transport to active travel results in a short-run increase in reported health satisfaction, driven predominantly by men. However, we do not observe significant shorter-run effects from switches from active travel to public transport.

In the longer-run at time $t + 2$ we find positive and significant effects on health satisfaction from a mode change from public to active travel driven mainly by women, and a corresponding decrease in health satisfaction for changes in the opposite direction (for men only). Mode changes from car to active travel at time t do not affect health at time $t + 2$. It is worth recalling that the health outcomes are self-reported and may be subject to differential reporting behavior (King, Murray, J., & Tandon, 2004). That is, individuals with similar levels of underlying health may report these differently due to perceptions of what may be regarded as healthy and preferences over health compared to other attributes. As long as such reporting behavior is time-invariant it does not present an issue for the analysis. However, it is possible that individuals adapt to changes in health over time, such that we observe initial health changes due to a change in commuting mode in the short to medium term but this then dissipates as time progresses and individuals become accustomed to their changed health. This might be the reason we observe less effects at $t + 2$ than at time $t + 1$. Oddly, while a mode change from active to car travel leads to a lowering of physical health in the longer run, we also see an improvement in mental health, particularly for women.

7 | SENSITIVITY CHECKS

7.1 | Seasonality in mode choices

It is possible that individuals may change their choice of commuting mode depending on weather conditions. Progressing into summer, individuals may increasingly opt to switch to active travel. Conversely individuals are more likely to switch to car or public transport in winter months. To control for seasonal effects we include the lag of the month of interview in our balancing and regression model. Results are reported in Table A4 and are consistent with the main results. Again, we observe an increase in mental health for both men and women when they switch from car to active travel and a decline in physical health for both groups when they switch from active to car. Similarly, the transition from public to active transport increases health satisfaction for men while the reverse transition decreases their mental health significantly, as previously observed.

As a further step, we divide the sample by seasons. Due to small sample sizes we combine spring/summer (typically warm and dry) and autumn/winter (typically cold and wet). These results are presented in Tables A5 and A6 and show that in summer and spring, mental health and satisfaction with own health for women increases when they switch from car to active modes of travel. Similarly, for men we observe an increase in physical health and satisfaction with own health when they change from public to active travel. Furthermore, we observe a decrease in mental health for men along with a decrease in physical health for women when they switch from active to public travel. In the winter months, we observe an increase in mental health for men when they switch from car to active modes of transport and a decline in physical health when they change from public to car travel. The only effect that we observe for women shows a decrease in their physical health when they switch from active to car travel. Broadly, in spring and summer months we observe more positive benefits of switching to active forms of travel, whilst in autumn and winter months we observe more negative health effects of switching from active to non-active forms of transport.

7.2 | Constant household location and job

So far, our estimation sample consists of individuals who do not change household address but we placed no restriction on their job characteristics. However, changes in commuting mode can also occur if individuals change jobs leading to a greater distance to travel. In further analysis, we select a subsample of individuals who report no change in household location or job characteristics. These estimates are reported in Table A7. Once again, these effects confirm our main results, although each of these effects are of a slightly higher magnitude compared to those in Table 6. Again, the main effects are observed for mens' health satisfaction which increases when they change from public to active travel and a significant decline in their mental health when they switch from active to public transport. We observe a decrease in physical health for women when they switch from active to car travel and an increase in physical health (at lower levels of significance) when they move from car to active transport.

7.3 | Panel attrition

As with all analyses that rely on panel data, results may be sensitive to attrition bias due to non-random drop out from the survey. Recalling that, for the main analysis, we include respondents observed across a minimum of three waves ($t-1$, t and $t+1$) where in wave t we observe a change to travel mode. The point at which the change to travel mode takes place can be in any particular wave (t) of the panel as long as the individual is also observed both in the previous ($t-1$) and subsequent ($t+1$) waves. When we consider a longer run effect, we observe outcomes at wave $t+2$. Accordingly, the sample of respondents are observed across four waves ($t-1$ to $t+2$). Due to attrition from the survey, the sample sizes for the sub-sets of respondents with observed outcomes at $t+2$ is smaller than for the sub-sets observed at wave $t+1$. This can be seen from a comparison of N in the results presented in Table 6 and those presented in Table A3. This change in sub-sample size allows us to investigate the likely role that survey attrition bias may play. When we compare the set of variables used in the EB and subsequent regression analysis across the two subsets of respondents, we observe the summary statistics presented in Table A8. As can be seen the means and standard deviations are very close. This provides prima-facie evidence for no significant attrition bias.

8 | CONCLUDING REMARKS

This paper evaluates the impact of a change in mode of commute on health. There is evidence on the gains to health from active modes of travel. Therefore, schemes to encourage active travel in the form of walking or cycling are being adopted by countries around the world. The majority of this evidence relies on (often dated) cross-sectional data and thus does not examine the effect of *changes* in travel mode on health. Of those few studies that do explore the effect of changes in mode, Martin et al. (2014) use fixed effects regressions to address the potential for selection bias. We improve on identification by employing an empirical strategy that combines matching techniques together with regression based analysis, to provide new evidence on the effect of commuting mode change on health. The approach has the advantage of being "doubly robust" to either poor matching but correct specification of the regression model, or complete covariate balance and misspecification of the regression model. Identification is, however, conditional on the validity of selection on observables. While we mitigate against failure of this assumption by matching on a wealth of pre-treatment characteristics including mode and health, some caution should be applied in interpreting the results due to the possibility of selection on unobservable characteristics influencing outcomes.

Using rich data taken from the UKHLS covering 2009–2016, we compare health outcomes (at various time periods) for individuals in employment who never change mode throughout the survey, with those who experience a mode change. Our main results indicate a significant increase in physical and mental health for commuters switching from car to active forms of transport, particularly for women. We further observe a decline in physical health for individuals of both sexes who switch from active travel to car. A change in mode from active travel to public transport leads to a decrease in reported mental health, largely for men, but we do not observe significant decreases in physical health. Mode changes in the opposite direction from public transport to active travel are associated with increases in reported satisfaction with health. The lack of an effect on physical health when changing between active and public transport may be due to accessing public transport requiring exercise, via walking to or from a bus or train station. As this is not the case for switches to and from car travel to active travel the benefits to physical health are more pronounced. Mode changes between car and public transport do not lead to notable affects on physical or mental health outcomes or satisfaction with health. Overall, our results lend support to UK policy initiatives designed to encourage people to move away from car commuting towards more active forms of travel. As well as the health effects estimated here, this will also help the UK government to meet its targets for reducing emissions.

ACKNOWLEDGMENTS

We are grateful for funding from the Economic and Social Research Council, grant reference: ES/P007546/1. Luke Munford acknowledges financial support from a MRC Skills Development Fellowship (MR/N015126/1). We thank participants, particularly Alister McGuire and Mireia Jofre-Bonet, of the Health Economists' Study Group meeting held in York, January, 2019, for helpful comments. The study uses data from Understanding Society (UK Household Longitudinal Study) which is available by application from the UK Data Service (https://ukdataservice.ac.uk/). The authors are not permitted to release the copy of UKHLS used for the study.

CONFLICT OF INTEREST

None of the authors have a conflict of interest with respect to this manuscript

ORCID

Nikita Jacob https://orcid.org/0000-0001-5546-4521
Luke Munford https://orcid.org/0000-0003-4540-6744
Nigel Rice https://orcid.org/0000-0003-0312-823X
Jennifer Roberts https://orcid.org/0000-0003-2883-7251

ENDNOTES

1 https://www.gov.uk/government/news/government-publishes-12-billion-plan-to-increase-cycling-and-walking

2 One possible approach is to use instrumental variables that predict changes to commuting mode and only affect health outcomes through this channel and not directly. Such instruments are difficult to find.

3 That all observable characteristics that determine both a change in commuting mode and health outcomes are included, and that conditional on these there are no unobservable characteristics that also influence mode and health outcomes.

4 Flint et al. (2014); Flint and Cummins (2015, 2016); Martin et al. (2014); Humphreys, Goodman, and Ogilvie (2013); Yang, Panter, Griffin, and Ogilvie (2012); Ettema and Smajic (2015).

5 Wener and Evans (2011); Bellet, Roman, and Kostis (1969); Ferenchak and Katirai (2015); Gatersleben and Uzzell (2007); Künn-Nelen (2015); Rissel, Petrunoff, Wen, and Crane (2014).

6 Friman et al. (2013); Gatersleben and Uzzell (2007); Páez and Whalen (2010); Turcotte (2005); Eriksson et al. (2013).

7 Some studies have exploited exogenous shocks to commuting through employer induced changes to work location (Jacob et al., 2019; van Ommeren and Gutierrez-i-Puigarnau, 2011). While we do not rule out changes in commuting due to a change in location of a workplace, we do not explicitly rely on this.

8 These are measured at time $t-1$ to avoid being contaminated by the treatment occurring at time t.

9 This is not treatment in a strict sense. We do not know the histories of individuals prior to changing modes. It maybe that those from the treated or control group may have switched modes in the past. However, we only consider individuals from the time they are first observed in the sample.

10 An alternative to matching is simply to regress the outcomes on the treatment indicator and the set of confounding variables. However, deriving causal effects with this approach is highly model dependent.

11 Alternatives include searching for natural experiments where a change in commuting model is exogenously imposed on individuals, but these are hard to find.

12 These inform of different patterns of attrition. For example, consider the following two individuals A is observed across waves 2 to 6, and B is observed in waves 2, 3, 5, 6, 7. Both are observed in five waves but have different sequences of observations leading to differences in the number of consecutive waves.

13 Note that where perfect balancing is achieved $\beta_x = 0$.

14 Observations are dropped if and when a subsequent mode change is observed.

15 The question on commuting at a given wave is in the present tense: *"How do you usually get to your place of work?"*. It can therefore be assumed that a change of mode took place at some time between waves $t-1$ and t. Accordingly, outcomes at time t can be considered short-run effects.

16 In the raw UKHLS, this variable is recorded on a 7 point scale, however, we recode it by combining responses 2–3 and 5–6.

17 $N = 31{,}736$ for analysis of outcomes at time t, consequently, $N = 18{,}156$ for outcomes at $t + 1$ and $N = 13{,}888$ at $t + 2$ (see Table 1).

18 Ideally we would control for the time spent on physical exercise by individuals to derive the true effects of switching to or from active modes of travel. However, this is not observed in our data.

REFERENCES

Abadie, A., & Imbens, G. (2011). Bias corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics*, *29*(1), 1–11.

Anable, J., & Gatersleben, B. (2005). All work and no play? The role of instrumental and affective factors in work and leisure journeys by different travel modes. *Transportation Research Part A: Policy and Practice*, *39*(2–3), 163–181.

Bellet, S., Roman, L., & Kostis, J. (1969). The effect of automobile driving on catecholamine and adrenocortical excretion. *American Journal of Cardiology*, *24*, 365–368.

Department for Transport. (2017). *Average trip time* by *trip purpose: England* (Technical report). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632823/nts0406.ods

Eriksson, L., Friman, M., & Gärling, T. (2013). Perceived attributes of bus and car mediating satisfaction with the work commute. *Transportation Research Part A: Policy and Practice*, *47*, 87–96.

Ettema, D., & Smajic, I. (2015). Walking, places and wellbeing. *The Geographic Journal*, *181*, 102–109.

Ferenchak, N., & Katirai, M. (2015). Commute mode and mental health in major metropolitan areas. *Transportation Letters*, *7*, 92–103.

Flint, E., & Cummins, S. (2015). Does active commuting protect against obesity in mid-life? Cross-sectional, observational evidence from UK biobank. *The Lancet*, *386*, S8.

Flint, E., & Cummins, S. (2016). Active commuting and obesity in mid-life: Cross-sectional, observational evidence from uk biobank. *Lancet Diabetes-Endocrinology*, *4*, 420–435.

Flint, E., Cummins, S., & Sacker, A. (2014). Associations between active commuting, body fat, and body mass index: Population based, cross sectional study in the United Kingdom. *The BMJ*, *349*, g4887.

Foster, M. (2009). *What are equivalence scales*, Paris, France: OECD Project on Income Distribution and Poverty.

Frank, L., Andresen, M., & Schmid, T. (2004). Obesity relationships with community design, physical activity and time spent in cars. *American Journal of Preventive Medicine*, *27*, 87–96.

Friman, M., Fujii, S., Ettema, D., Grling, T., & Olsson, L. (2013). Psychometric analysis of the satisfaction with travel scale. *Transportation Research Part A*, *48*, 132–145.

Gatersleben, B., & Uzzell, D. (2007). Affective appraisals of the daily commute: Comparing perceptions of drivers, cyclists, walkers and users of public transport. *Environment and Behavior*, *39*, 416–431.

Genter, J., Donovan, S., Petrenas, B., & Badland, H. (2008). *Valuing the health benefits of active transport modes*, New Zealand: New Zealand Transport Agency Research Report (359).

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, *80*(2), 223–255.

Hainmueller, J. (2012). Entropy balancing for causal effects: A mutivariate reweighting methods to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25–46.

Hainmueller, J., & Xu, Y. (2013). ebalance: A stata package for entropy balancing. *Journal of Statistical Software*, *54*(7), 1–18.

Hamer, M., & Chida, Y. (2008). Active commuting and cardiovascular risk: A meta-analytic review. *Preventive Medicine*, *46*(1).

Heckman, J., & Robb, R. (1985). Alternative models for evaluating the impact of interventions: An overview. *Journal of Econometrics*, *30*, 239–267.

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*, 199–236.

Humphreys, D., Goodman, A., & Ogilvie, D. (2013). Associations between active commuting and physical and mental wellbeing. *Preventive Medicine*, *57*, 135–139.

Jacob, N., Munford, L., Rice, N., & Roberts, J. (2019). The disutility of commuting? The effect of gender and local labour markets. *Regional Science and Urban Economics*, *77*, 264–275.

Jones, A., Koolman, X., & Rice, N. (2006). Health-related non-response in the british household panel survey and european community household panel: Using inverse-probability-weighted estimators in non-linear models. *Journal of the Royal Statistical Society (Series A)*, *169*(3), 543–569.

King, G., Murray, C., J., S., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, *98*, 191–207.

Künn-Nelen, A. (2015). Does commuting affect health? *Health Economics*, *25*, 984–1004.

Lancee, S., Veenhoven, R., & Burger, M. (2017). Mood during commute in The Netherlands: What way of travel feels best for what kind of people? *Transportation Research Part A: Policy and Practice*, *104*, 195–208.

Laverty, A., Mindell, J., Webb, E., & Millett, C. (2013). Active travel to work and cardiovascular risk factors in the United Kingdom. *American Journal of Preventive Medicine*, *45*(3), 282–288.

Litman, T. (2010). *Evaluating public transportation health benefits*, Victoria, Canada: Victoria Transport Policy Institute. Retrieved from http://vtpi.org/tran_health.pdf

MacDonald, J., Stokes, R., Cohen, D., Kofner, A., & Ridgeway, G. (2010). The effect of light rail transit on body mass index and physical activity. *American Journal of Preventive Medicine*, *39*, 105–112.

Martin, A., Goryakin, Y., & Suhrcke, M. (2014). Does active commuting improve psychological wellbeing? Longitudinal evidence from eighteen waves of the british household panel survey. *Preventive Medicine*, *69*, 296–303.

Martin, A., Panter, J., Suhrcke, M., & Ogilvie, D. (2015). Impact of changes in mode of travel to work on changes in body mass index: Evidence from the british household panel survey. *Journal of Epidemiology and Community Health*, *69*(8), 753–761.

Morris, E. (2015). Should we all just stay home? Travel, out-of-home activities, and life satisfaction. *Transportation Research Part A: Policy and Practice*, *78*, 519–536.

NICE. (2012). *National institute for health and care excellence. Walking and cycling: Local measures to promote walking and cycling as forms of travel or recreation (public health guidance 41)*, London, UK: NICE. Retrieved from www.nice.org.uk/guidance/ph41

Páez, A., & Whalen, K. (2010). Enjoyment of commute: A comparison of different transportation modes. *Transportation Research Part A: Policy and Practice*, *44*(7), 537–549.

Rabl, A., & deNazelle, A. (2012). Benefits of shift from car to active transport. *Transport Policy*, *19*, 121–131.

Rissel, C., Petrunoff, N., Wen, L., & Crane, M. (2014). Travel to work and self-reported stress: Findings from work places survey in south west Sydney. *Journal of Transport and Health*, *1*, 50–53.

Roberts, J., Hodgson, R., & Dolan, P. (2011). It's driving her mad: Gender differences in the effects of commuting on psychological health. *Journal of Health Economics*, *30*(5), 1064–1076.

Scheepers, C., Wendel-Vos, G., den Broeder, J., van Kempen, E., van Wesemael, P., & Schuit, A. (2014). Shifting from car to active transport: A systematic review of the effectiveness of interventions. *Transportation Research Part A: Policy and Practise*, *70*, 264–280.

Stokols, D., Novaco, R., Stokols, J., & Campbell, J. (1978). Traffic congestion, type a behavior and stress. *Journal of Applied Psychology*, *63*, 467–480.

Turcotte, M. (2005). Like commuting? Workers perceptions of their daily commute, *Canadian Social Trends*. *Statistics Canada Catalogue No. 11-008*, *82*, 35–40.

Turcotte, M. (2011). Commuting to work: Results of the 2010 general social survey. *Canadian Social Trends*. *Statistics Canada Catalogue No. 11-008-X*, *92*, 25–36.

van Ommeren, J., & Gutierrez-i-Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, *41*(1), 1–8.

Verbeek, M., & Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, *33*, 681–703.

Ware, J. E., Keller, S. D., & Kosinski, M. (2002). *SF-12: How to score the SF-12 physical and mental health summary scales*, Boston, US: New England Medical Center, Health Institute.

Wener, R., & Evans, G. (2007). A morning stroll: Levels of physical activity in car and mass transit commuting. *Environment and Behavior*, *39*(1), 62–74.

Wener, R., & Evans, G. (2011). Comparing stress of car and train commuters. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*, 111–116.

Yang, L., Panter, J., Griffin, S., & Ogilvie, D. (2012). Associations between active commuting and physical activity in working adults: Cross-sectional results from the commuting and health in cambridge study. *Preventive Medicine*, *55*, 453–457.

Ziebarth, N. (2010). Measurement of health, health inequality and repoting heterogeneity. *Social Science & Medicine*, *71*, 116–124.

How to cite this article: Jacob N, Munford L, Rice N, Roberts J. Does commuting mode choice impact health? *Health Economics*. 2020;1–24. https://doi.org/10.1002/hec.4184

APPENDIX A

TABLE A1 Number of mode changes in the analytical sample

	Number		Percent	
	Observations	Individuals	Observations	Individuals
Criteria	***NT***	***N***	***NT***	***N***
Full analytic sample	106,195	31,736	100	100
#. of changes				
0	82,213	26,177	77%	82%
1	14,193	3654	13%	12%
2	7168	1452	7%	5%
3	1936	344	2%	1%
4	572	92	1%	0%
5	106	16	0%	0%
6	7	1	0%	0%

TABLE A2 Entropy balancing by gender for outcomes at time t

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health.Sat	Women (8) Health.Sat	Men (9) Health.Sat
Car->Public	−0.209	−0.466	0.183	0.054	0.137	0.008	−0.007	−0.011	−0.003
	(0.280)	(0.394)	(0.381)	(0.316)	(0.435)	(0.460)	(0.046)	(0.061)	(0.068)
N	43,832	24,788	19,044	43,832	24,788	19,044	43,693	24,705	18,988
Treated	764	433	331	764	433	331	761	431	330
Control	43,068	24,355	18,713	43,068	24,355	18,713	42,959	24,291	18,668
Car->Active	0.099	0.214	0.001	0.656***	1.037***	−0.009	0.080**	0.074	0.083
	(0.225)	(0.310)	(0.325)	(0.243)	(0.347)	(0.338)	(0.035)	(0.048)	(0.053)
N	44,135	24,944	19,191	44,135	24,944	19,191	43,995	24,861	19,134
Treated	1067	589	478	1067	589	478	1061	586	475
Control	43,068	24,355	18,713	43,068	24,355	18,713	42,959	24,291	18,668
Public->Car	0.356	0.462	0.276	0.014	0.188	−0.162	0.063	0.100	0.014
	(0.280)	(0.402)	(0.394)	(0.350)	(0.525)	(0.480)	(0.048)	(0.065)	(0.071)
N	6263	3380	2883	6263	3380	2883	6227	3356	2871
Treated	853	486	367	853	486	367	851	484	367
Control	5410	2894	2516	5410	2894	2516	5383	2877	2506
Public->Active	0.286	0.402	0.073	−0.042	−0.157	0.131	0.148**	0.040	0.290***
	(0.374)	(0.454)	(0.626)	(0.464)	(0.641)	(0.674)	(0.066)	(0.086)	(0.102)
N	5836	3144	2692	5836	3144	2692	5801	3121	2680
Treated	426	250	176	426	250	176	425	249	176
Control	5410	2894	2516	5410	2894	2516	5383	2877	2506
Active->Car	−0.424*	−0.021	−0.870**	0.351	0.240	0.465	−0.069*	−0.049	−0.106*
	(0.257)	(0.367)	(0.380)	(0.293)	(0.404)	(0.420)	(0.040)	(0.053)	(0.061)
N	6385	3946	2439	6385	3946	2439	6359	3931	2428
Treated	1056	598	458	1056	598	458	1050	593	457
Control	5329	3348	1981	5329	3348	1981	5316	3340	1976
Active->Public	0.350	0.610	−0.078	0.212	0.372	−0.054	−0.018	−0.011	−0.025
	(0.364)	(0.525)	(0.509)	(0.520)	(0.776)	(0.656)	(0.072)	(0.098)	(0.106)
N	5734	3583	2151	5734	3583	2151	5714	3572	2142
Treated	405	235	170	405	235	170	404	234	170
Control	5329	3348	1981	5329	3348	1981	5316	3340	1976

Note: Individuals present for at least two waves. Dependent variables measured at t, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

TABLE A3 Entropy balancing by gender for outcomes at time $t + 2$

	SF12 Indicators at $t + 2$						Other Indicators at $t + 2$		
	Overall	Women	Men	Overall	Women	Men	Overall	Women	Men
Variables	(1) SF12PCS	(2) SF12PCS	(3) SF12PCS	(4) SF12MCS	(5) SF12MCS	(6) SF12MCS	(7) Health.Sat	(8) Health.Sat	(9) Health.Sat
Car->Public	−0.409	−1.173	0.729	−0.385	−0.639	−0.349	−0.043	−0.132	0.082
	(0.511)	(0.768)	(0.595)	(0.559)	(0.809)	(0.739)	(0.073)	(0.093)	(0.112)
N	19,095	10,994	8101	19,095	10,994	8101	19,026	10,949	8077
Treated	521	297	224	521	297	224	519	295	224
Control	34,922	20,024	14,898	34,922	20,024	14,898	34,850	19,981	14,869
Car->Active	0.292	0.170	0.391	0.126	0.097	0.071	0.041	0.072	−0.006
	(0.325)	(0.466)	(0.436)	(0.389)	(0.539)	(0.545)	(0.053)	(0.072)	(0.076)
N	19,262	11,079	8183	19,262	11,079	8183	19,192	11,034	8158
Treated	738	407	331	738	407	331	735	405	330
Control	34,922	20,024	14,898	34,922	20,024	14,898	34,850	19,981	14,869
Public->Car	0.258	0.942	−0.621	0.751	−0.986	0.735	0.091	0.165	0.036
	(0.470)	(0.640)	(0.673)	(0.755)	(0.759)	(0.754)	(0.077)	(0.101)	(0.121)
N	2381	1258	1123	1123	1258	1123	2367	1249	1118
Treated	535	314	221	535	314	221	534	313	221
Control	3894	2023	1871	3894	2023	1871	3880	2015	1865
Public->Active	0.118	0.364	0.101	−0.389	0.536	−1.433	0.240**	0.329**	0.084
	(0.638)	(0.825)	(0.982)	(0.848)	(0.984)	(1.544)	(0.108)	(0.142)	(0.171)
N	2194	1143	1051	2194	1143	1051	2180	1134	1046
Treated	253	152	101	253	152	101	252	151	101
Control	3894	2023	1871	3894	2023	1871	3880	2015	1865
Active->Car	−0.959**	−0.777	−0.895	0.838*	1.506**	0.018	−0.108*	−0.065	−0.098
	(0.418)	(0.556)	(0.637)	(0.476)	(0.681)	(0.722)	(0.063)	(0.081)	(0.101)
N	2564	1580	984	2564	1580	984	2554	1575	979
Treated	679	380	299	679	380	299	675	377	298
Control	3931	2471	1460	3931	2471	1460	3925	2469	1456
Active->Public	−0.976	−0.640	−1.606*	0.356	0.192	0.785	−0.152	−0.013	−0.318**
	(0.674)	(0.949)	(0.942)	(0.821)	(1.267)	(1.032)	(0.111)	(0.152)	(0.162)
N	2268	1417	851	2268	1417	851	2261	1414	847
Treated	245	143	102	245	143	102	244	142	102
Control	3931	2471	1460	3931	2471	1460	3925	2469	1456

Note: Individuals present for at least 4 waves. Dependent variables measured at $t + 2$, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

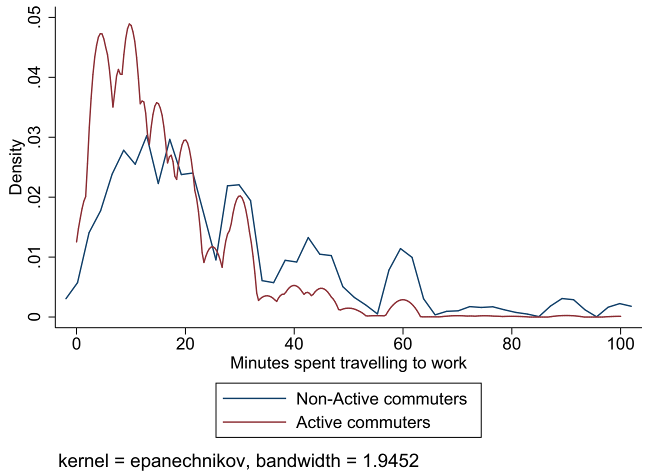

FIGURE A1 Commuting time for active versus non-active travel

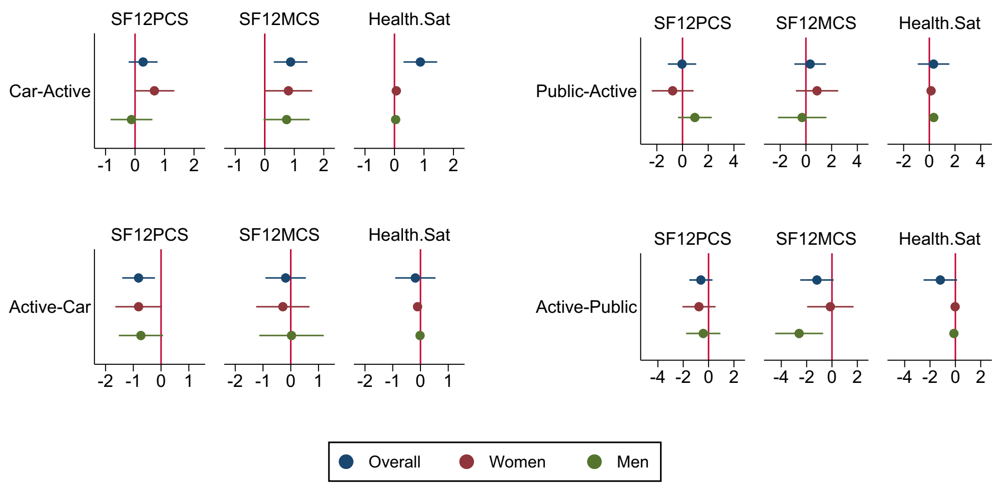

FIGURE A2 Effect of mode changes on health at $t + 1$

TABLE A4 Robustness check 2: with lag of month, Outcome at $t + 1$

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health. Sat	Women (8) Health. Sat	Men (9) Health. Sat
Car->Public	0.022	−0.020	0.070	0.010	−0.116	0.157	0.034	0.001	0.096
	(0.351)	(0.494)	(0.495)	(0.390)	(0.537)	(0.561)	(0.056)	(0.075)	(0.085)
N	29,714	16,954	12,760	29,714	16,954	12,760	29,617	16,895	12,722
Treated	646	369	277	646	369	277	644	367	277
Control	39,636	22,518	17,118	39,636	22,518	17,118	39,544	22,463	17,081
Car->Active	0.279	0.541	−0.101	0.880***	0.818**	0.739*	0.050	0.060	0.037
	(0.250)	(0.357)	(0.361)	(0.289)	(0.411)	(0.394)	(0.042)	(0.056)	(0.062)
N	29,937	17,064	12,873	29,937	17,064	12,873	29,839	17,005	12,834

(Continues)

TABLE A4 (Continued)

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health. Sat	Women (8) Health. Sat	Men (9) Health. Sat
Treated	909	498	411	909	498	411	906	496	410
Control	39,636	22,518	17,118	39,636	22,518	17,118	39,544	22,463	17,081
Public->Car	−0.074	0.397	−0.734	0.344	0.713	0.098	0.024	0.045	0.020
	(0.358)	(0.516)	(0.481)	(0.436)	(0.668)	(0.618)	(0.059)	(0.077)	(0.091)
N	3909	2094	1815	3909	2094	1815	3889	2083	1806
Treated	707	412	295	707	412	295	706	411	295
Control	4639	2473	2166	4639	2473	2166	4619	2461	2158
Public->Active	−0.022	−0.760	0.970	0.280	0.567	−0.302	0.230***	0.144	0.356***
	(0.562)	(0.829)	(0.684)	(0.629)	(0.839)	(0.955)	(0.083)	(0.109)	(0.127)
N	3609	1912	1697	3609	1912	1697	3589	1901	1688
Treated	330	188	142	330	188	142	329	187	142
Control	4639	2473	2166	4639	2473	2166	4619	2461	2158
Active->Car	−0.805***	−0.781*	−0.802*	−0.180	−0.368	−0.000	−0.076	−0.095	−0.013
	(0.301)	(0.430)	(0.411)	(0.369)	(0.486)	(0.593)	(0.048)	(0.065)	(0.075)
N	4098	2542	1556	4098	2542	1556	4083	2533	1550
Treated	861	487	374	861	487	374	856	483	373
Control	4688	2964	1724	4688	2964	1724	4678	2959	1719
Active->Public	−0.693	−0.922	−0.419	−1.017	0.060	−2.575***	−0.062	−0.033	−0.112
	(0.471)	(0.675)	(0.699)	(0.673)	(0.929)	(0.979)	(0.084)	(0.118)	(0.121)
N	3670	2305	1365	3670	2305	1365	3659	2299	1360
Treated	333	196	137	333	196	137	332	195	137
Control	4688	2964	1724	4688	2964	1724	4678	2959	1719

Note: Individuals present for at least three waves. Dependent variables measured at $t+1$, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and lags of wave and month of interview (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

TABLE A5 Seasonality effect I: Spring and Summer, Outcome at $t+1$

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health.Sat	Women (8) Health.Sat	Men (9) Health.Sat
Car-Public	−0.140	−0.237	−0.136	0.201	0.429	0.176	0.079	−0.024	0.221*
	(0.508)	(0.681)	(0.775)	(0.550)	(0.671)	(0.858)	(0.082)	(0.111)	(0.121)
N	14,965	8560	6405	14,965	8560	6405	14,926	8536	6390
Treated	313	175	138	313	175	138	312	174	138
Control	19,978	11,350	8628	19,978	11,350	8628	19,941	11,329	8612

TABLE A5 (Continued)

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health.Sat	Women (8) Health.Sat	Men (9) Health.Sat
Car-Active	0.195	0.298	0.146	1.256***	1.953***	0.301	0.135**	0.182**	0.126
	(0.376)	(0.549)	(0.521)	(0.396)	(0.572)	(0.531)	(0.057)	(0.077)	(0.084)
N	15,102	8626	6476	15,102	8626	6476	15,063	8602	6461
Treated	480	260	220	480	260	220	479	259	220
Control	19,978	11,350	8628	19,978	11,350	8628	19,941	11,329	8612
Public-Car	0.556	1.089	0.013	0.393	1.111	−0.297	0.177**	0.166	0.223*
	(0.490)	(0.695)	(0.686)	(0.565)	(0.746)	(0.812)	(0.086)	(0.117)	(0.128)
N	1950	1035	915	1950	1035	915	1940	1029	911
Treated	350	203	147	350	203	147	349	202	147
Control	2323	1229	1094	2323	1229	1094	2315	1224	1091
Public-Active	1.328*	0.558	2.271**	−0.017	0.373	−1.273	0.324***	0.202	0.479***
	(0.735)	(1.074)	(0.984)	(0.894)	(1.272)	(1.263)	(0.118)	(0.166)	(0.159)
N	1806	942	864	1806	942	864	1797	937	860
Treated	166	88	78	166	88	78	166	88	78
Control	2323	1229	1094	2323	1229	1094	2315	1224	1091
Active-Car	−0.467	−0.404	−0.524	0.269	0.255	0.215	−0.100	−0.117	−0.034
	(0.419)	(0.611)	(0.534)	(0.507)	(0.686)	(0.771)	(0.067)	(0.088)	(0.103)
N	2090	1301	789	2090	1301	789	2084	1297	787
Treated	419	234	185	419	234	185	418	233	185
Control	2412	1520	892	2412	1520	892	2407	1517	890
Active-Public	−0.861	−1.944**	−0.046	−0.758	1.598	−3.520**	−0.043	−0.049	0.004
	(0.639)	(0.952)	(0.779)	(0.955)	(1.210)	(1.397)	(0.124)	(0.167)	(0.181)
N	1880	1188	692	1880	1188	692	1875	1185	690
Treated	166	96	70	166	96	70	166	96	70
Control	2412	1520	892	2412	1520	892	2407	1517	890

Note: Individuals present for at least 3 waves. Dependent variables measured at $t + 1$, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

TABLE A6 Seasonality effect II: Autumn and winter, outcome at $t + 1$

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health.Sat	Women (8) Health.Sat	Men (9) Health.Sat
Car-Public	0.108	−0.106	0.257	−0.113	−0.361	0.212	0.002	−0.007	0.020
	(0.494)	(0.731)	(0.625)	(0.556)	(0.799)	(0.709)	(0.079)	(0.106)	(0.115)
N	14,749	8394	6355	14,749	8394	6355	14,691	8359	6332
Treated	333	194	139	333	194	139	332	193	139
Control	19,658	11,168	8490	19,658	11,168	8490	19,603	11,134	8469
Car-Active	0.388	0.858*	−0.251	0.593	0.092	1.087**	0.013	0.057	−0.028
	(0.334)	(0.448)	(0.490)	(0.406)	(0.579)	(0.552)	(0.062)	(0.082)	(0.095)
N	14,835	8438	6397	14,835	8438	6397	14,776	8403	6373
Treated	429	238	191	429	238	191	427	237	190
Control	19,658	11,168	8490	19,658	11,168	8490	19,603	11,134	8469
Public-Car	−0.891*	−0.533	−1.645**	0.347	−0.298	1.112	−0.120	−0.066	−0.198
	(0.511)	(0.776)	(0.659)	(0.589)	(0.829)	(0.815)	(0.078)	(0.101)	(0.122)
N	1959	1059	900	1959	1059	900	1949	1054	895
Treated	357	209	148	357	209	148	357	209	148
Control	2316	1244	1072	2316	1244	1072	2304	1237	1067
Public-Active	−1.265	−2.013	−0.135	0.761	0.967	0.635	0.157	0.108	0.266
	(0.840)	(1.257)	(0.797)	(0.848)	(0.937)	(1.452)	(0.118)	(0.141)	(0.175)
N	1803	970	833	1803	970	833	1792	964	828
Treated	164	100	64	164	100	64	163	99	64
Control	2316	1244	1072	2316	1244	1072	2304	1237	1067
Active-Car	−1.253***	−1.441**	−0.900	−0.583	−0.532	−0.470	−0.013	0.010	0.020
	(0.436)	(0.629)	(0.579)	(0.517)	(0.644)	(0.853)	(0.067)	(0.089)	(0.103)
N	2008	1241	767	2008	1241	767	1999	1236	763
Treated	442	253	189	442	253	189	438	250	188
Control	2276	1444	832	2276	1444	832	2271	1442	829
Active-Public	−0.416	−0.158	−0.730	−1.690*	−1.640	−2.035	−0.086	0.083	−0.229
	(0.691)	(0.917)	(1.114)	(0.948)	(1.321)	(1.367)	(0.105)	(0.145)	(0.159)
N	1789	1117	672	1789	1117	672	1783	1114	669
Treated	166	100	66	166	100	66	165	99	66
Control	2276	1444	832	2276	1444	832	2271	1442	829

Note: Individuals present for at least 3 waves. Dependent variables measured at $t + 1$, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

TABLE A7 Robustness check 3: Constant Household and Job, outcome at $t+1$

Variables	Overall (1) SF12PCS	Women (2) SF12PCS	Men (3) SF12PCS	Overall (4) SF12MCS	Women (5) SF12MCS	Men (6) SF12MCS	Overall (7) Health. Sat	Women (8) Health. Sat	Men (9) Health. Sat
Car->Public	−0.088	0.260	−0.533	0.554	0.219	1.036*	−0.038	−0.032	−0.042
	(0.446)	(0.613)	(0.634)	(0.456)	(0.653)	(0.593)	(0.070)	(0.090)	(0.110)
N	23,423	13,404	10,019	23,423	13,404	10,019	23,338	13,352	9986
Treated	448	256	192	448	256	192	445	254	191
Control	31,953	18,228	13,725	31,953	18,228	13,725	31,872	18,180	13,692
Car->Active	0.570*	0.695	0.431	0.380	0.487	0.073	0.027	0.024	0.026
	(0.293)	(0.428)	(0.396)	(0.347)	(0.506)	(0.466)	(0.049)	(0.067)	(0.071)
N	23,596	13,481	10,115	23,596	13,481	10,115	23,512	13,429	10,083
Treated	670	358	312	670	358	312	668	356	312
Control	31,953	18,228	13,725	31,953	18,228	13,725	31,872	18,180	13,692
Public->Car	−0.467	−0.122	−0.888	0.339	1.203	−0.450	−0.040	−0.014	−0.073
	(0.453)	(0.662)	(0.568)	(0.544)	(0.826)	(0.780)	(0.070)	(0.094)	(0.107)
N	2991	1613	1378	2991	1613	1378	2973	1601	1372
Treated	478	287	191	478	287	191	477	286	191
Control	3680	1960	1720	3680	1960	1720	3662	1948	1714
Public->Active	−0.397	−0.915	0.146	0.477	1.169	−0.807	0.269**	0.212	0.334**
	(0.725)	(1.119)	(0.771)	(0.806)	(1.013)	(1.304)	(0.112)	(0.146)	(0.164)
N	2782	1480	1302	2782	1480	1302	2764	1468	1296
Treated	219	121	98	219	121	98	218	120	98
Control	3680	1960	1720	3680	1960	1720	3662	1948	1714
Active->Car	−0.918**	−1.210**	−0.811	−0.640	−0.402	−0.938	−0.094*	−0.067	−0.112
	(0.370)	(0.506)	(0.506)	(0.459)	(0.586)	(0.746)	(0.056)	(0.072)	(0.091)
N	3443	2172	1271	3443	2172	1271	3430	2164	1266
Treated	620	362	258	620	362	258	616	359	257
Control	4084	2597	1487	4084	2597	1487	4075	2592	1483
Active->Public	−1.074*	−1.072	−1.213	−1.453*	0.236	−3.629***	−0.214**	−0.213	−0.152
	(0.645)	(0.981)	(0.992)	(0.855)	(1.238)	(1.308)	(0.108)	(0.152)	(0.158)
N	3097	1968	1129	3097	1968	1129	3088	1963	1125
Treated	202	122	80	202	122	80	202	122	80
Control	4084	2597	1487	4084	2597	1487	4075	2592	1483

Note: Individuals present for at least three waves. Household location and job characteristics held constant. Dependent variables measured at $t+1$, and are increasing in good health. Controls matched to treated using entropy balancing at $t-1$, prior to regression of outcomes on treatment (at t), conditioning on covariates and wave dummies (at $t-1$). We also balance on attrition variables, consecutive waves and number of waves. Covariates include age, number of kids, job hours, marital status, household income, commuting time and initial health. Estimates for Health Satisfaction are coefficients from an ordered probit model. Standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

TABLE A8 Sample attrition at $t + 1$ and $t + 2$

	Sample $t + 1$			Sample $t + 2$		
Variable	***N***	**Mean**	**SD**	***N***	**Mean**	**SD**
SF12PCS	86,519	52.94	8.01	73,715	52.97	8.00
SF12MCS	86,519	49.94	8.80	73,715	49.96	8.75
Satisfaction with own health	86,350	3.48	1.05	73,581	3.48	1.04
Male	86,519	0.44	0.50	73,715	0.44	0.50
Age	86,519	43.13	10.92	73,715	43.66	10.46
University level qualification	74,892	0.46	0.50	63,812	0.46	0.50
College level qualification	74,892	0.21	0.40	63,812	0.20	0.40
School level qualification	74,892	0.20	0.40	63,812	0.20	0.40
Household size	86,519	3.06	1.30	73,715	3.05	1.29
Number of children	86,519	0.72	0.98	73,715	0.72	0.97
Married/Cohabiting	86,410	0.74	0.44	73,623	0.75	0.43
Usual hours worked	86,519	33.49	9.92	73,715	33.61	9.71
Log household income	86,368	7.57	0.53	73,594	7.58	0.52

Note: Summary statistics for sample with outcomes at time $t + 1$ and at time $t + 2$.